Statistics of mass substructure from strong gravitational lensing: quantifying the mass fraction and mass function

Simona Vegetti* & L. V. E. Koopmans
Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, the Netherlands

ABSTRACT

A Bayesian statistical formalism is developed to quantify the level at which the mass-function ($dN/dm \propto m^{-\alpha}$) and the projected cumulative mass fraction (f) of (CDM) substructure in strong gravitational-lens galaxies, with arcs or Einstein rings, can be recovered as function of the lens-survey parameters and the detection threshold of the substructure mass. The method is applied to different sets of mock data to explore a range of observational limits: (i) the number of lens galaxies in the survey, (ii) the mass threshold, M_{low}, for the detection of substructures and (iii) the uncertainty of the measured substructure masses. We explore two different priors on the mass function slope: a uniform prior and a Gaussian prior with $\alpha = 1.90 \pm 0.1$. With a substructure detection threshold $M_{low} = 3 \times 10^8 M_\odot$, the number of lenses available now ($n_l = 30$), a true dark-matter mass fraction in (CDM) substructure $\leqslant 1.0\%$ and a prior of $\alpha = 1.90 \pm 0.1$, we find that the upper limit of f can be constrained down to a level $\leqslant 1.0\%$ (95% CL). In the case of a uniform prior the complete substructure mass distribution (i.e. mass fraction and slope) can only be characterized in a number of favourable cases with a large number of detected substructures. This can be achieved by an increase of the resolution and the signal-to-noise ratio of the lensed images. In the case of a Gaussian prior on α, instead, it is always possible to set stringent constraints on both parameters. We also find that lowering the detection threshold has the largest impact on the ability to recover α, because of the (expected) steep mass-function slope. In the future, thanks to new surveys with telescopes, such as SKA, LSST and JDEM and follow-up telescopes with high-fidelity data, a significant increase in the number of known lenses (i.e. $\gg 10^4$) will allow us to recover the satellite population in its completeness. For example, a sample of 200 lenses, equivalent in data-quality to the *Sloan Lens ACS Survey* and a detection threshold of $10^8 M_\odot$, allows one to determine $f = 0.5 \pm 0.1\%$ (68% CL) and $\alpha = 1.90 \pm 0.2$ (68% CL).

Key words: gravitational lensing — dark matter — galaxies: structure — galaxies: haloes — methods: statistics

1 INTRODUCTION

In the context of the cold-dark-matter paradigm, a significant number of substructures, with a steep mass function, is expected to populate the dark halo of galaxies. In galaxies as massive as the Milky Way, for example, of the order of 10^4 substructures are predicted inside the virial radius (Diemand et al. 2008; Springel et al. 2008), although only about 20 have been so far observed (Zucker et al. 2004; Willman et al. 2005; Belokurov et al. 2006; Grillmair 2006; Martin et al. 2006; Sakamoto & Hasegawa 2006; Zucker et al. 2006a,b; Belokurov et al. 2007; Ibata et al. 2007; Irwin et al. 2007; Majewski et al. 2007; Walsh et al. 2007; Zucker et al. 2007; Belokurov et al. 2008). A clear comparison between the simulated and the physical reality, however, is strongly hampered by the difficulty of directly observing substructures in

distant galaxies, as well as in the Local Group.

While major improvements in the observations and numerical simulations have not yet significantly alleviated the satellite crisis, new techniques have been proposed for the indirect and direct detection of subhaloes that may have a high mass-to-light ratio. In our own Galaxy, CDM substructures can be, in principle, identified via their effect on stellar streams (Ibata et al. 2002; Mayer et al. 2002) or via the dark matter annihilation signal from their high-density centres (Bergström et al. 1999; Calcáneo-Roldán & Moore 2000; Stoehr et al. 2003; Colafrancesco et al. 2006); gravitational lensing, on the other hand, allows for direct detection (measurement of the substructure gravitational signature) in the central regions of galaxies through flux-ratio anomalies and distortions of extended Einstein rings and arcs. (e.g. Mao & Schneider 1998; Metcalf & Madau 2001; Dalal & Kochanek 2002; Koopmans 2005). Interestingly enough, results based on flux-ratio anomalies reverse the satellite crisis with a recovered mass fraction in sub-

* E-mail: vegetti@astro.rug.nl

structure which seems to be larger than predicted by numerical simulations (e.g. Mao et al. 2004; Macciò & Miranda 2006). While this discrepancy might not be easily accommodated by an increase in the resolution of simulations, the correct interpretation of the flux-ratio anomalies is still subject of discussion.

In Vegetti & Koopmans (2009), we introduced a new adaptive-grid method, based on a Bayesian analysis of the surface brightness distribution of highly magnified Einstein rings and arcs, that allows the identification and precise quantification of substructure in single lens galaxies. This technique does not depend on the nature of dark matter, on the shape of the main galaxy halo, strongly on the density profile of the substructure, nor on the dynamical state of the system. It can be applied to local galaxies as well as to high redshift ones, as long as the lensed images are highly magnified, extended and have a high signal-to-noise ratio. Unlike flux ratio anomalies, extended optical images are little affected by differential scattering in the radio or microlensing in the optical and X-ray. If substructures are located close to the lensed images, the method allows the determination of both their mass and position, although as the distance between the substructure and the Einstein ring increases, the mass model becomes more degenerate. Finally, thanks to its Bayesian framework, the method of Vegetti & Koopmans (2009) requires hardly any subjective intervention into the modelling and any assumption can be objectively tested through the Bayesian Evidence (MacKay 1992).

In this paper, we show how results from data sets with more than one lens system (i.e. the number of detections and their masses) can be combined to statistically constrain the fraction of dark matter in substructure and their mass function, as function of the survey parameters and limits. The combination of multiple data sets becomes important when trying to constrain the slope of the substructure mass function. In a single lens potential, in fact, the number of detectable substructures can, in certain cases, be as small as zero or one. More than one lens is, therefore, required in order to improve statistics and properly *sample* the mass function. On the contrary an upper limit to the mass fraction in substructure can always be set.

The paper is organized as follows. In Section 2, we outline the statistical formalism of satellite detection; in Section 3, application of the method is discussed and in Section 4 conclusions are drawn.

2   BAYESIAN INTERPRETATION OF SUBSTRUCTURE DETECTIONS

In this section, we outline a Bayesian formalism that allows us to give a statistical interpretation to the detection of dark substructure in gravitational lens galaxies and to recover the properties of the substructure population. Thanks to Bayes' theorem, the likelihood of measuring a mass substructure, can be translated into probability density distributions for the substructure mass fraction f and their mass function $dN/dm \propto m^{-\alpha}$ (when dN/dm is normalised to unity, we refer to it as dP/dm), as function of the mass measurement errors and the model parameters. The latter include the minimum and maximum mass of dark matter substructure, $M_{\min}$ and $M_{\max}$, between which the mass fraction is defined, and the lowest and highest mass we can detect, M_{low} and M_{high}. More precisely, M_{low} should not be interpreted as a hard detection threshold, but as a statistical limit above which we believe a detection to be significant, at the level set by the mass measurement error σ_m (i.e. with a signal-to-noise ratio of M_{low}/σ_m).

2.1   Likelihood of the substructure measurements

We derive an expression for the Likelihood of observing n_s substructures for a given sample of n_l lenses. We assume the cumulative dark-matter mass fraction of substructure $f(< R/R_{\text{vir}})$, within a cylinder of projected radius R of the lens, to be the same for all lens galaxies and the detections of multiple substructures in one galaxy to be independent from one another. We also assume that the number of substructures populating a given galactic halo fluctuates with a distribution which is Poissonian. We know that not all these substructures are observable, but only those with the right combination of mass and position.

The Likelihood of measuring n_s substructures, each of mass m_i, in a single galaxy, in general can be expressed as the probability density of having n_s substructures in the considered lens, times the normalised probability density $P(m_i, R \mid \boldsymbol{p}, \alpha)$ of actually observing the mass m_i within the projected radius R, where $\boldsymbol{p} = \left\{ M_{\min}, M_{\max}, M_{\text{low}}, M_{\text{high}} \right\}$ is a vector containing all the fixed model parameters introduced above, so that

$$\mathcal{L}(n_s, \boldsymbol{m} \mid \alpha, f, \boldsymbol{p}) = \frac{e^{-\mu(\alpha, f, <R)} \mu(\alpha, f, < R)^{n_s}}{n_s!} \prod_{i=1}^{n_s} P(m_i, R \mid \boldsymbol{p}, \alpha) , \tag{1}$$

where $\boldsymbol{m}$ contains all the substructure masses m_i and $\mu(\alpha, f, < R)$ is the expectation value of the number of substructures in a generic aperture with dark matter mass $M_{\text{DM}}(< R)$, which will be discussed in more details in the following section. The second factor in equation (1) describes the likelihood with which a substructure can be identified as a function of its mass m_i and position R. Although the specific shape of $P(m_i, R \mid \boldsymbol{p}, \alpha)$ might change from one system to another, its overall trend is essentially the same for all lenses: high mass perturbations, on or close to the lensed images are the most likely to be confidently observed. For each considered lens system, $P(m_i, R \mid \boldsymbol{p}, \alpha)$ can be reconstructed by Monte Carlo explorations of the model parameter space (Vegetti & Koopmans 2009).

Here, for the sake of simplicity, we assume that the probability of measuring a mass m_i for a single substructure is a Gaussian function independent of position. This is indeed true for small regions around the Einstein ring of the lensed images, so that $P(m_i, R \mid \boldsymbol{p}, \alpha)$ reduces to

$$P(m_i \mid \boldsymbol{p}, \alpha) = \frac{\int_{M_{\min}}^{M_{\max}} \frac{dP}{dm}\big|_{true} \frac{e^{-(m-m_i)^2/2\sigma_m^2}}{\sqrt{2\pi}\sigma_m} dm}{\int_{M_{\text{low}}}^{M_{\text{high}}} \int_{M_{\min}}^{M_{\max}} \frac{dP}{dm}\big|_{true} \frac{e^{-(m-m')^2/2\sigma_m^2}}{\sqrt{2\pi}\sigma_m} dm \, dm'} , \tag{2}$$

where the Gaussian convolution expresses the scatter in the substructure mass due to a measurement uncertainty and $dP/dm|_{true}$ is the true substructure mass function as defined in equation (6). The assumption of Gaussian errors on the substructure mass is motivated by results from Vegetti & Koopmans (2009).

Equation (1) can be easily extended to the case in which more than one lens is considered. Because the identification of a substructure in one system does not influence what we infer about another satellite in another lens, the Likelihood function for a set of n_l lenses is simply the product of the independent likelihoods of individual detections,

$$\mathcal{L}(\{n_s, \boldsymbol{m}\} \mid \alpha, f, \boldsymbol{p}) = \prod_{k=1}^{n_l} \mathcal{L}(n_{s,k}, \boldsymbol{m}_k \mid \alpha, f, \boldsymbol{p}) . \tag{3}$$

We now go through the details of the expectation value $\mu(\alpha, f, < R)$, characterising the Poisson distribution of the number of substructures in the potential of a generic lens galaxy.

2.2 Substructure expectation value

In the ideal case of infinite sensitivity to a mass perturbation, one would be able to recover the full mass function between $M_{\min}$ and $M_{\max}$. In practice, only substructures with mass between M_{low} and M_{high} are observable, hence the expectation value for observable substructures is

$$\mu(\alpha, f, < R) = \mu_0(\alpha, f, < R, \boldsymbol{p}) \int_{M_{\text{low}}}^{M_{\text{high}}} \frac{dP}{dm}\bigg|_{true} dm \qquad (4)$$

with the expectation from the full mass function given by

$$\mu_0(\alpha, f, < R, \boldsymbol{p}) = \frac{f(< R) \, M_{\text{DM}}(< R)}{\int_{M_{\min}}^{M_{\max}} m \, \frac{dP}{dm}\big|_{true} \, dm} =$$

$$= f(< R) \, M_{\text{DM}}(< R) \begin{cases} \frac{(2-\alpha)\left(M_{\max}^{1-\alpha} - M_{\min}^{1-\alpha}\right)}{(1-\alpha)\left(M_{\max}^{2-\alpha} - M_{\min}^{2-\alpha}\right)} & \alpha \neq 2 \quad , \alpha \neq 1 \\[2ex] -\frac{\left(M_{\max}^{-1} - M_{\min}^{-1}\right)}{\log\left(M_{\max} / M_{\min}\right)} & \alpha = 2 \\[2ex] \frac{\log\left(M_{\max}/M_{\min}\right)}{\left(M_{\max} - M_{\min}\right)} & \alpha = 1 \end{cases}$$
$$(5)$$

Hence, we assume the normalised true mass function to be given by a power-law

$$\frac{dP}{dm}\bigg|_{true} = \begin{cases} \frac{(1-\alpha)\, m^{-\alpha}}{\left(M_{\max}^{1-\alpha} - M_{\min}^{1-\alpha}\right)} & \alpha \neq 1 \\[2ex] \frac{m^{-\alpha}}{\log\left(M_{\max}/M_{\min}\right)} & \alpha = 1 \end{cases} \qquad (6)$$

$M_{\text{DM}}(< R)$ and $f(< R)$ are the cumulative mass in dark matter and the cumulative fraction of dark matter in subhaloes within the considered radius, respectively. However, the presence of noise on the data and the statistical uncertainty with which masses are measured introduce a scatter in the observed mass function, so that detections can be spread inside or outside our observational limits. The significance of this effect depends on the substructure mass, with lower masses being affected by a larger relative uncertainty. The observed mass function $dP/dm|_{conv}$ can then be written as a convolution of the true mass function with the error distribution, which we assume to be Gaussian, hence

$$\mu(\alpha, f, < R, \boldsymbol{p}) = \mu_0(\alpha, f, < R, \boldsymbol{p}) \int_{M_{\text{low}}}^{M_{\text{high}}} \frac{dP}{dm}\bigg|_{conv} dm =$$

$$\mu_0(\alpha, f, < R, \boldsymbol{p}) \int_{M_{\text{low}}}^{M_{\text{high}}} \int_{M_{\min}}^{M_{\max}} \frac{dP}{dm}\bigg|_{true} \frac{e^{-(m-m')^2/2\sigma_m^2}}{\sqrt{2\pi}\sigma_m} dm \, dm' . \quad (7)$$

A mathematical proof for this procedure can be found in the Appendix.

2.3 Posterior probability function of α and f

Given a set of observations, in which a certain number of substructures are identified and their masses are quantified for one or more lenses, equations (1) and (3) can be used to infer the mass fraction and the mass function of the underlying subhalo population. Bayes' theorem relates the Likelihood function of the observations to the joint posterior probability of α and f in the following way

$$P(\alpha, f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p}) = \frac{\mathcal{L}(\{n_s, \boldsymbol{m}\} \mid \alpha, f, \boldsymbol{p}) \, P(\alpha, f \mid \boldsymbol{p})}{P(\{n_s, \boldsymbol{m}\} \mid \boldsymbol{p})}, \qquad (8)$$

where $P(\alpha, f \mid \boldsymbol{p})$ is the prior probability density distribution function of α and f. For the mass fraction we assume a non-informative uniform prior between the limits $f_{\min} = 0$ and $f_{\max} = 1$; while we test two different priors for α. We assume in one case a uniform distribution between $\alpha_{min} = 1.0$ and $\alpha_{max} = 3.0$ and in the other case a Gaussian function with centre on 1.90 and a standard deviation of 0.1, as found in almost all numerical simulations. We refer to the next section for a more detailed description of the prior probability density distributions.

2.4 Dark matter mass

As shown in equation (5), the number of substructures expected in a given potential is a function of the dark mass $M_{\text{DM}}(< R)$. In this section we show how this mass can be empirically estimated. Specifically, we are interested in the cumulative mass within a narrow annulus $\Delta R = 2 \, \delta R = 0.6''$ centred around the Einstein radius R_{E}, where the formalism introduced can be considered valid. In the approximation of a small annulus, the dark matter mass contained in it can be approximated as

$$M_{\text{DM}} \approx 4\pi \, R_{\text{E}} \, \Sigma_{\text{DM}}(R_{\text{E}}) \, \delta R = 4\pi \, R_{\text{E}} \left[1 - \frac{\Sigma_*}{\Sigma_{\text{tot}}}\right] \Sigma_{\text{tot}} \, \delta R . \qquad (9)$$

Σ_{tot} and Σ_* are the total and the stellar projected mass density, respectively. It has been found by many authors, that the total mass density has a profile which is close to isothermal (e.g Gerhard et al. 2001; Koopmans & Treu 2002; Koopmans et al. 2006; Czoske et al. 2008; Koopmans et al. 2009). Similarly, a Jaffe (1983) profile approximates the stellar mass distribution well, i.e.,

$$\Sigma_{\text{tot}}(R) = \frac{\Sigma_c \, R_{\text{E}}}{2R} \qquad (10)$$

and

$$\Sigma_*(R) = C\left\{\frac{1}{4\,\tilde{r}} + \frac{1}{2\pi}\left[\left(1 - \tilde{r}^2\right)^{-1} - \left(1 - \tilde{r}^2\right)^{-3/2}\left(2 - \tilde{r}^2\right)\cosh^{-1}(\tilde{r}^{-1})\right]\right\} \qquad (11)$$

with $\tilde{r} = R/r_s$.

Σ_c is the critical surface density for lensing and r_s is the scaling radius of the Jaffe profile, which relates to the effective radius of the lens galaxy as $r_s = R_e/0.74$. We have assumed in the above equations that the Einstein radius and the effective radius are related to each other by $R_e = 2R_{\text{E}}$, which is approximately the case for the average SLACS lens (Bolton et al. 2008). Obviously, this is not exactly true for any of these lenses, but we are not interested in analysing an exact reproduction of the SLACS sample but just an average realisation of it. This can also be considered as a fair realisation of a typical massive early-type galaxy, given that both the internal and environmental properties of the SLACS lenses do not significantly depart from those of other early-type galaxies with comparable velocity dispersion and baryonic properties (Bolton et al. 2006; Treu et al. 2009). This assumption is in any case justified and does not influence our results; in fact, only the cumulative dark matter mass that is probed by all lenses is of relevance and its average value is not altered by the assumption of $R_e = 2R_{\text{E}}$.

The normalization constant $C = 0.74\Sigma_c$ can be derived by imposing $\Sigma_{\text{tot}} \to \Sigma_* \approx \frac{Cr_s}{4R}$ for $R \to 0$, that is by imposing that asymptotically for $R \to 0$ the mass density becomes that of stars only, i.e. the *maximum bulge* assumption.

Because we assume that R_e/r_s is a constant and R_e/R_{E} is also a constant, on average the projected DM mass fraction within an annulus

of $2\delta R$ around the Einstein radius is a constant with a value of about 63%. Hence we find that on average $\Sigma_{DM} \approx 0.5 \times 0.63 \, \Sigma_c$ at the Einstein radius for all lenses. This fraction is consistent with other observations (e.g. Gavazzi et al. 2008; Schechter & Wambsganss 2002). The number of substructures thus becomes only a function of the size of the aperture and the critical density, which itself only weakly depends on the source and lens redshifts.

3   DATA REALISATION AND ANALYSIS

In this section, we use a series of mock data sets to show that the formalism presented here allows us to constrain, in different situations, the properties of the substructure population. Different data sets of lens galaxies with $n_l = 10, 30$ or 200 are analysed. A sample based on the masses and radii of the SLACS lenses, ranked from the highest to the lowest mass enclosed within R_E, is used to construct the dark matter mass function, while results from the method in Vegetti & Koopmans (2009) are used to set the observational limits and the substructure mass uncertainty. In each lens, the substructures are distributed in mass according to $dP/dm|_{true}$ with $M_{min} = 4.0 \times 10^6 M_\odot$ and $M_{max} = 4.0 \times 10^9 M_\odot$ (Diemand et al. 2007a,b) and fluctuate in number with a Poisson probability distribution of expectation value $\mu(\alpha, f, < R)$, given by equation (4).

Although it was shown by Vegetti & Koopmans (2009) that the detection probability is a joint function of the substructure mass and position, this probability can be assumed to be independent of the perturber's position if a sufficiently small annulus around the Einstein ring is considered, in which case the equations presented in the previous sections hold. We consider a region of ± 0.3 arcsec around R_E, over which a typical SLACS lens shows a reasonable surface brightness of its images and within which we might expect to detect CDM substructures using the method of Vegetti & Koopmans (2009). The extent of this area is, also, such that the mass fraction of substructures can be considered constant in radius, hence $f(R) = $ const. over ΔR.

Each mock data set is characterized by a different fraction f_{true} of substructures, while the true slope of the mass function is kept fixed at $\alpha_{true} = 1.90$, as suggested by numerical simulations. Results from the latest high-resolution numerical simulations seem to indicate a dark matter mass fraction in substructures within a cylinder of projected 10 kpc which is between 0.3%[1] (Mark Vogelsberger, private communication) and 0.5% (Diemand et al. 2008). We therefore discuss three different cases $f_{true} = 0.1, 0.5, 2.5\%$. The latter high fraction is included because it is close to that suggested by the median value inferred from flux-ratio anomaly studies (Dalal & Kochanek 2002).

3.1   Observational limits on the substructure mass

We explore the effect of different values of the lowest detectable mass M_{low}, which is set at the statistical threshold above which we are confident that other effects do not create too many false events. We have shown in Vegetti & Koopmans (2009) that, given the current HST data quality of the SLACS lenses, a lower mass limit for

---

[1] This value has been obtained by including only particles within $R_{200_{crit}}$, adding masses of all those particles that are in subhaloes within the considered cylinder and considering the median over 100 projected directions. No extrapolation beyond the resolution limit or cut on the subhalo mass is involved.

---

a significant detection can be set around $10^8 M_\odot$, depending on how close the perturbers are located with respect to the lensed images and the structure of the lensed images. However, these limits have been determined for cases not affected by systematic errors; we adopt $M_{low} = 0.3, 1.0, 3.0 \times 10^8 M_\odot$. We set a finite upper mass limit $M_{high} = M_{max}$.

3.2   Priors on α and f

While the mass fraction of satellites is the most uncertain parameter, most studies seem to agree on the mass function, with values of the slope α ranging from 1.8 to 2.0 (e.g. Helmi et al. 2002; Gao et al. 2004; Diemand et al. 2008; Springel et al. 2008). This is a direct consequence of the assumed cold nature of the dark matter particles. We analyse two scenarios: the first, relying on results from numerical simulations, assumes for α a Gaussian prior centred at α_{true} with standard deviation $\sigma_\alpha = 0.1$; the second scenario, allowing for more freedom, has a uniform prior between 1.0 and 3.0. In general, the first case can be seen as a test of N-body simulations and the second as a test of nature itself, although in the specific case of this paper the data have been created with a combination of fraction and slope typical of a standard cosmology. The former prior, obviously, provides a well-defined mass-function slope, but it also reduces the uncertainty in the mass fraction, which is less well-defined than α and can vary considerably between similar simulations. Different phenomena can affect the mass fraction, as for example the resolution of the simulations, or the lack, in high resolution simulations, of gas physics, that could sensibly influence the substructure survival (i.e. f could be even higher than what the simulations suggest in the inner regions of the galactic haloes). We assume, conservatively, for f a uniform prior ranging between 0% and 100%.

3.3   Results in the limit of no mass measurement errors

We present results for cases in which the errors on the mass measurements can be neglected. Specifically, this translates into convolving the mass distribution not with a Gaussian but with a delta function around m_i in the equations of Section 2.

In Fig. 1 we show the joint probability contours $P(\alpha, f \mid \{n_s\}, \boldsymbol{m}\}, \boldsymbol{p})$ and the marginalized probabilities $P(f \mid \{n_s\}, \boldsymbol{m}\}, \boldsymbol{p})$ and $P(\alpha \mid \{n_s\}, \boldsymbol{m}\}, \boldsymbol{p})$, for systems containing 10 randomly realised lenses. Specifically the plotted contours contain, in the limit of a Gaussian distribution, respectively, 68%, 95% and 97.2% of the marginalized probability function.

In the case of a uniform prior, while a good upper limit to f can always be set, little can be said about the slope, which can only be constrained for a limited number of favourable physical and observational conditions, such as $(f = 0.5\%, M_{low} = 0.3 \times 10^8 M_\odot)$, $(f = 2.5\%, M_{low} = 0.3 \times 10^8 M_\odot)$ and $(f = 2.5\%, M_{low} = 1.0 \times 10^8 M_\odot)$. In Fig. 2 an equivalent plot is presented for systems with $n_l = 30$; although even more stringent limits can be given for f, we are still unable to recover the underlying mass function for most of the possible scenarios.

The situation can be substantially improved by increasing the number of detectable substructures with a larger number of lenses. To provide insight into future capabilities, results from three samples of 200 lenses with $f = 0.5\%$ and respectively $M_{low} = 0.3 \times 10^8 M_\odot$, $M_{low} = 1.0 \times 10^8 M_\odot$, and $M_{low} = 3.0 \times 10^8 M_\odot$ are given in Fig. 3. Currently, no uniform sample with 200 lenses with high signal-to-noise ratio and high resolution (equivalent to that of

HST) is available. However, forthcoming ground and space-based instruments (e.g LSST/JDEM, EVLA, e-Merlin, LOFAR and SKA) can provide these numbers (in fact beyond these) in the coming 5–10 years and the required data quality by dedicated follow-up. A detailed characterisation of the CDM mass-function, through the technique of Vegetti & Koopmans (2009), could therefore be realisable in the coming years if investments are made in large high-resolution and high-sensitivity lens surveys with these instruments (see Koopmans et al. 2009). As can be seen from both Figs. 1 and 2, in the case of a Gaussian prior on α, tight limits can be set on the mass fraction for all possible combinations of the considered parameters.

The results from this section are summarised in Table 1, where we report the input values for each parameter, the recovered maximum-posterior values $(f_{MP}; \alpha_{MP})$ at which $P(\alpha, f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ reaches its maximum, the median, the 68% and 95% (σ_{68} and σ_{95}), confidence levels of the marginalized probabilities $P(f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ and $P(\alpha \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ for both the cases of a uniform and a Gaussian prior on α. σ_{68}, and σ_{95}, in the particular case of a Gaussian distribution, respectively represent the 1σ and 2σ error.

3.4 The effect of mass measurement errors

We explore now how the presence of uncertainty in the mass measurements affects our analysis. In particular, we consider three cases: $\sigma_m = 0.1 \times 10^8 M_\odot$ with $M_{low} = 0.3 \times 10^8 M_\odot$, $\sigma_m = 1/3 \times 10^8 M_\odot$ with $M_{low} = 10^8 M_\odot$, and $\sigma_m = 1.0 \times 10^8 M_\odot$ with $M_{low} = 3.0 \times 10^8 M_\odot$ i.e. a limiting signal-to-noise ratio of $M_{low}/\sigma_m = 3$. The lens systems analysed here have $n_l = 10, 30, 200$ lenses and a mass fraction in substructures $f_{true} = 0.5\%$. Relative likelihood contours are plotted in the three panels of Fig. 4. These have to be compared with the equivalent *no-error* results in Figs. 1, 2 and 3. Results are reported in Table 2.

The effect of measurement errors on the substructure mass depends on the form of prior adopted for α, with uniform priors being more strongly affected than Gaussian ones. Errors as large as $\sigma_m = 1/3 \times 10^8 M_\odot$ combined with mass threshold of $M_{low} = 10^8 M_\odot$, can significantly influence even systems with 200 lenses . Specifically, these systems were created by drawing masses between M_{min} and M_{max}, then scattering each mass with a Gaussian distribution (i.e. mimicking the measurement uncertainty) and then using only those objects that fall within the detection range $\left[M_{low}, M_{high}\right]$ to constrain the fraction and the mass distribution. We refer to the Appendix for a mathematical proof that this way of proceeding is equivalent to drawing between M_{low} and M_{high} with a Poisson probability density distribution of expectation values given by equation (4).

4 CONCLUSIONS

We have introduced a statistical formalism for the interpretation and the generalisation of subhalo detection in gravitational lens galaxies, that allows us to quantify the mass fraction and the mass function of CDM substructures. Given mock sets of lenses, with properties typical of a CDM cosmology, we have analysed how well the true parameters can be recovered. The formalism depends on several parameters, such as e.g. the number of lenses, the mass detection threshold and the measurement errors. It has a very general nature and, in principle, it could be used to statistically analyse substructure detection by flux ratio anomalies or timedelay/astrometric perturbations as well. In practice, these methods

would first need to show that their mass estimates are meaningful and second they would have to determine the probability distribution of flux-ratio anomalies or perturbations, either as function of the lens geometry or marginalized over all model parameters, which could be rather computationally expensive. The method has been tested on several mock data sets, with parameter settings based on our knowledge of the SLACS lenses. Several physical and observational scenarios have been considered. We list here the main results:

• If the number of arc/ring lens systems is $\ll 100$, as is the case for current surveys (e.g. SLACS), the ability to constrain the mass fraction and the mass function of satellites still depends on the form of prior which is assumed for α. In particular, if results from numerical simulations are assumed to hold and a Gaussian prior with $\alpha_{true} = 1.9 \pm 0.1$ is adopted, we are able to constrain both α and f for any data sets containing a number of lenses $n_l \geqslant 10$, with improved limits for either increasing mass fractions, decreasing detection threshold or increasing number of lenses. If instead a wider range of possibilities is explored by assuming a uniform prior, one can still set strong limits on f, even for values as low as $f = 0.1\%$ and a detection threshold $M_{low} = 0.3 \times 10^8 M_\odot$, but the mass function slope can be recovered only in a limited number of favourable cases, characterized by high mass fraction and low detection threshold.

• Our ability to constrain α could be considerably improved either by increasing our sensitivity to substructures, i.e. by increasing the quality of the data, or by increasing the number of analysed objects. Although competing with the quality of HST seems at the moment difficult, future surveys such as LSST/JDEM in the optical and EVLA, e-MERLIN, LOFAR and SKA in the radio, will surely lead to an increase in the number of known lenses by several orders of magnitude (see Koopmans et al. 2009; Marshall et al. 2009) and dedicated optical and/or radio follow-up could provide equivalent or better data quality than HST. We expect therefore, in the foreseeable future to be able to characterise the galactic subhalo population with stringent constraint, both on the mass fraction and slope.

• Although we have not explicitly performed a model comparison between different cosmologies, as for example CDM versus Warm Dark Matter (this would require an extra marginalization of the parameter space), the formalism introduced here, combined with the sensitivity of our method to CDM substructures, will allows us, in the future, to discriminate among these two scenarios and thus test the physics of dark matter.
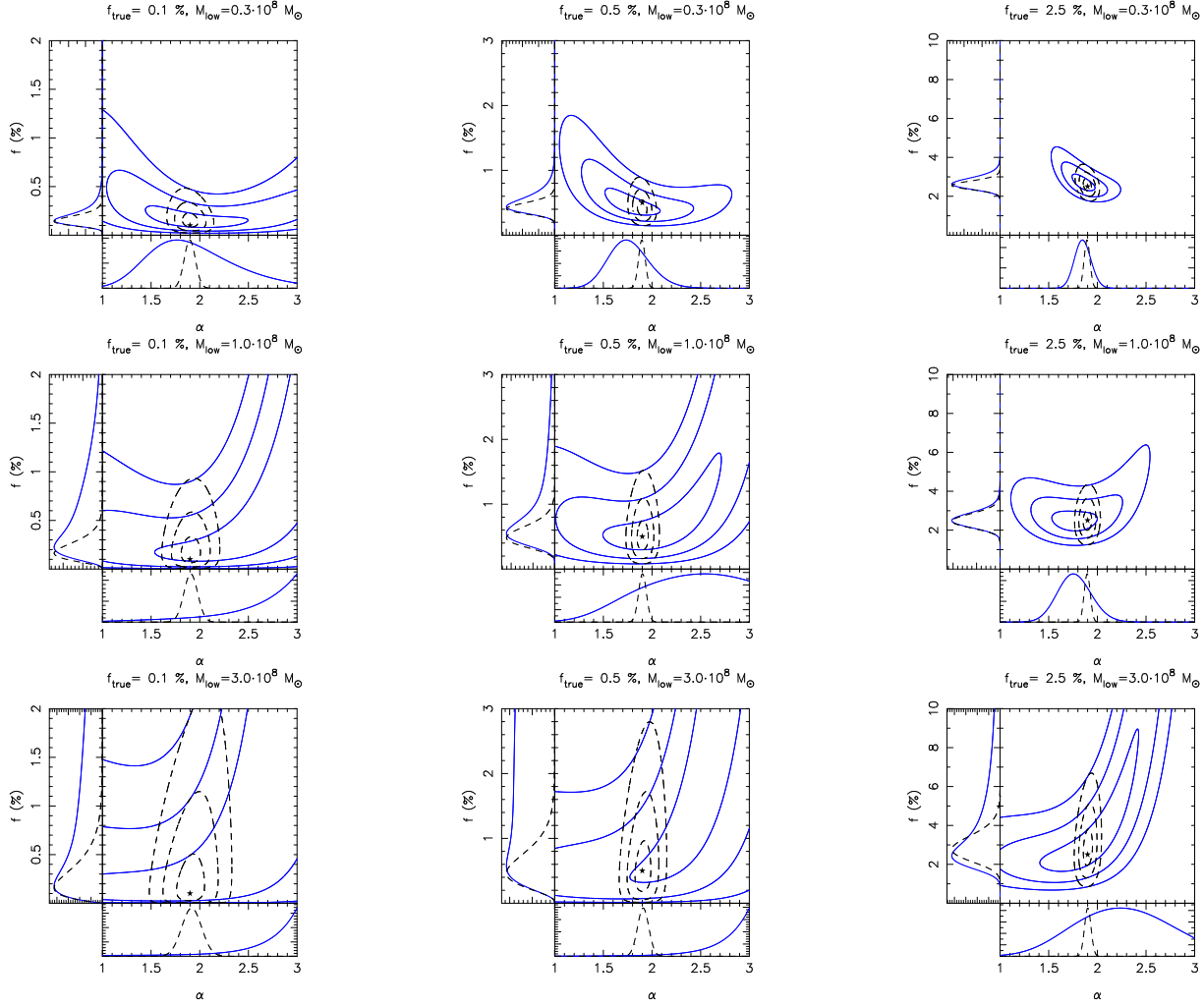
Figure 1. Results for systems with 10 randomly realised lenses. In each panel, the joint probability $P(\alpha, f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ contours and marginalized probabilities $P(f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ and $P(\alpha \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ are given for a uniform prior (solid lines) and for a Gaussian prior on α (dashed lines). Moving from one panel to next the substructure fraction f increases from left to right and the detection limit M_{low} increases from top to bottom.

References

Belokurov V., Walker M. G., Evans N. W., et al., 2008, ApJL, 686, L83

Belokurov V., Zucker D. B., Evans N. W., et al., 2006, ApJL, 647, L111

Belokurov V., Zucker D. B., Evans N. W., et al., 2007, ApJ, 654, 897

Bergström L., Edsjö J., Gondolo P., Ullio P., 1999, Phys. Rev. D, 59, 043506

Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703

Bolton A. S., Treu T., Koopmans L. V. E., Gavazzi R., Moustakas L. A., Burles S., Schlegel D. J., Wayth R., 2008, ApJ, 684, 248

Calcáneo-Roldán C., Moore B., 2000, Phys. Rev. D, 62, 123005

Colafrancesco S., Profumo S., Ullio P., 2006, A&A, 455, 21

Czoske O., Barnabè M., Koopmans L. V. E., Treu T., Bolton A. S., 2008, MNRAS, 384, 987

Dalal N., Kochanek C. S., 2002, ApJ, 572, 25

Diemand J., Kuhlen M., Madau P., 2007a, ApJ, 667, 859

Diemand J., Kuhlen M., Madau P., 2007b, ApJ, 657, 262

Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, Nature, 454, 735

Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819

Gavazzi R., Treu T., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., Marshall P. J., 2008, ApJ, 677, 1046

Gerhard O., Kronawitter A., Saglia R. P., Bender R., 2001, AJ, 121, 1936

Grillmair C. J., 2006, ApJL, 645, L37

Helmi A., White S. D., Springel V., 2002, Phys. Rev. D, 66, 063502

Ibata R., Martin N. F., Irwin M., Chapman S., Ferguson A. M. N., Lewis G. F., McConnachie A. W., 2007, ApJ, 671, 1591

Ibata R. A., Lewis G. F., Irwin M. J., Quinn T., 2002, MNRAS, 332, 915

Irwin M. J., Belokurov V., Evans N. W., et al., 2007, ApJL, 656, L13

Jaffe W., 1983, MNRAS, 202, 995

Koopmans L. V. E., 2005, MNRAS, 363, 1136

Koopmans L. V. E., Bolton A., Treu T., et al., 2009, ApJL, 703, L51
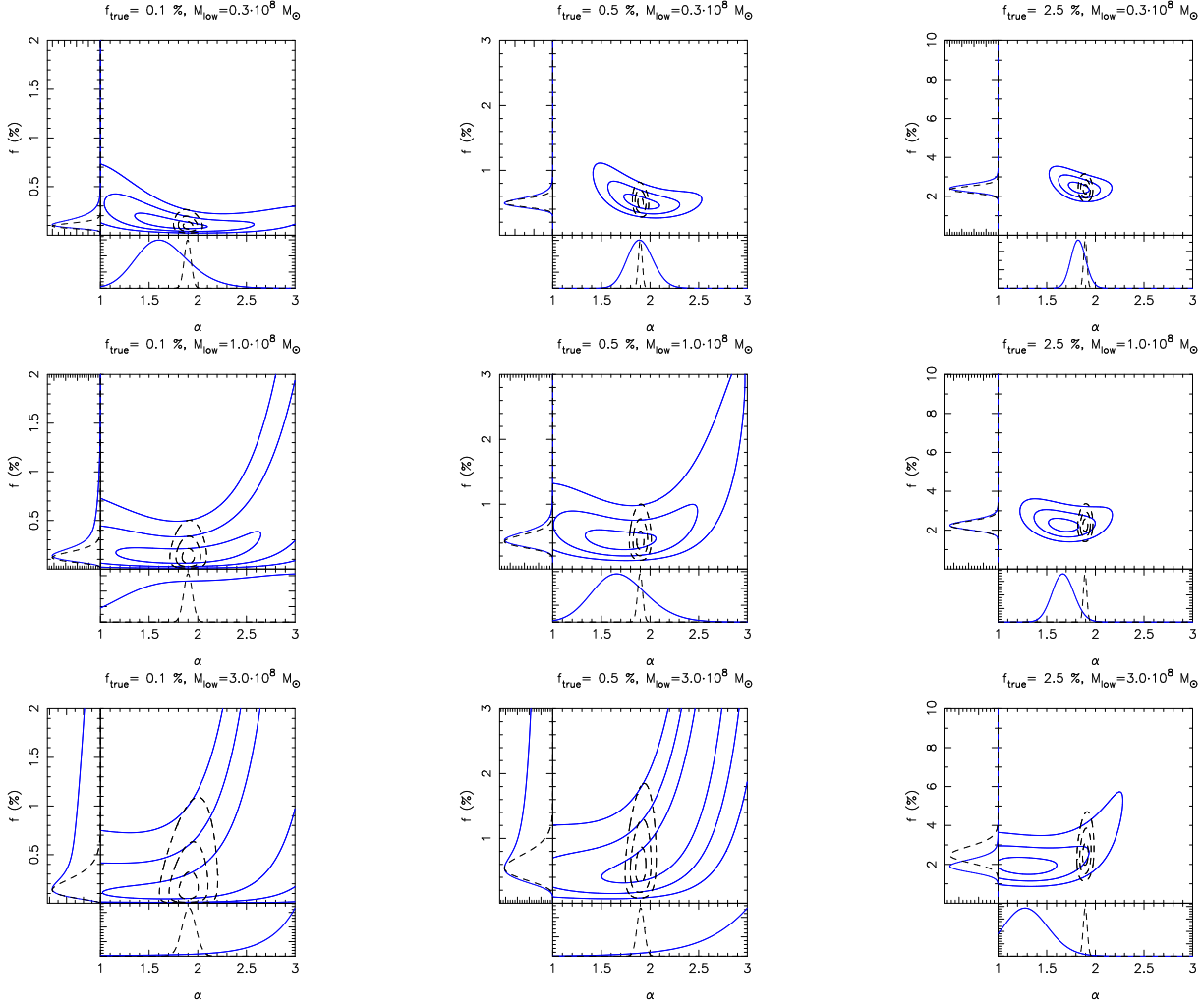
Figure 2. Similar to Fig. 1 for systems with 30 lenses.
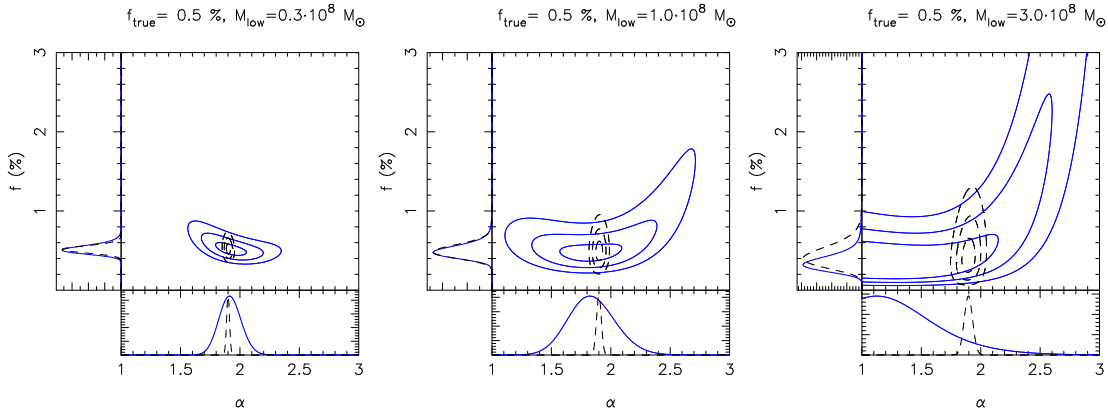
Figure 3. Results for three samples with 200 randomly realised lenses with $f = 0.5\%$ and respectively $M_{\rm low} = 0.3 \times 10^8 M_\odot$ (left panel), $M_{\rm low} = 1.0 \times 10^8 M_\odot$ (middle panel) and $M_{\rm low} = 3.0 \times 10^8 M_\odot$ (right panel). The joint probability $P(\alpha, f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ contours and marginalized probabilities $P(f \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ and $P(\alpha \mid \{n_s, \boldsymbol{m}\}, \boldsymbol{p})$ for a uniform prior (solid lines) and for a Gaussian prior in α (dashed lines) are shown.
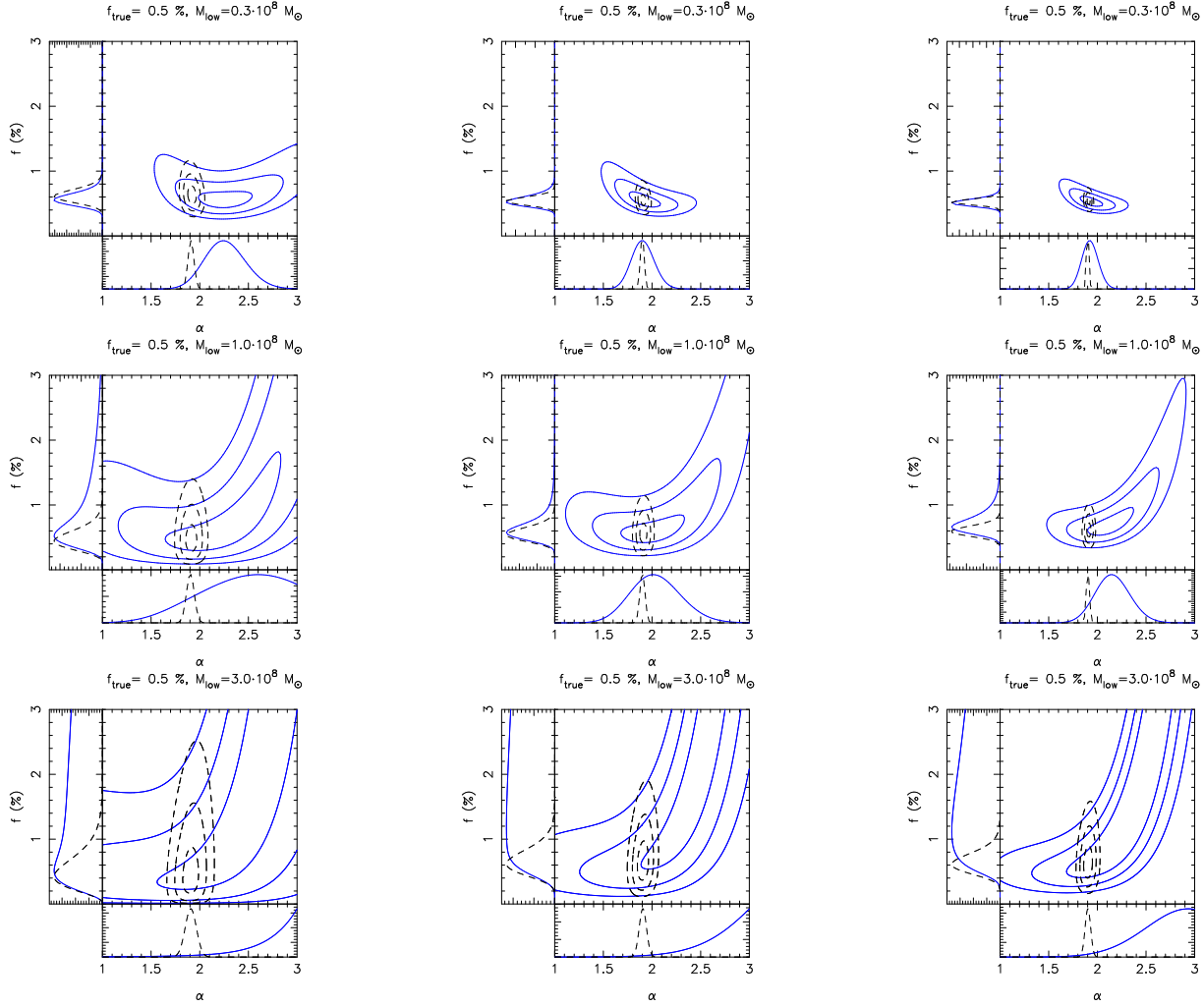
Figure 4. Effect of different measurement error levels on the substructure mass. Similar as Fig. 1 for systems with $f = 0.5\%$, $M_{\text{low}} = 0.3 \times 10^8 M_\odot$ and $\sigma_m = 0.1 \times 10^8 M_\odot$ (upper panels), with $M_{\text{low}} = 10^8 M_\odot$ and $\sigma_m = 1/3 \times 10^8 M_\odot$ (middle panels) and with $M_{\text{low}} = 3 \times 10^8 M_\odot$ and $\sigma_m = 1.0 \times 10^8 M_\odot$ (lower panels). Results for 10, 30 and 200 lenses are plotted in the left, middle and right panels, respectively.

Koopmans L. V. E., et al., 2009, arXiv:0902.3186v2

Koopmans L. V. E., Treu T., 2002, ApJL, 568, L5

Koopmans L. V. E., Treu T., Bolton A. S., Burles S., Moustakas L. A., 2006, ApJ, 649, 599

Macciò A. V., Miranda M., 2006, MNRAS, 368, 599

MacKay D. J. C., 1992, PhD thesis, Caltech

Majewski S. R., Beaton R. L., Patterson R. J., et al., 2007, ApJL, 670, L9

Mao S., Jing Y., Ostriker J. P., Weller J., 2004, ApJ, 604, L5

Mao S., Schneider P., 1998, MNRAS, 295, 587

Marshall P. J., Auger M., Bartlett J. G., et al., 2009, arXiv:0902.2963, 2010, 194

Martin N. F., Ibata R. A., Irwin M. J., Chapman S., Lewis G. F., Ferguson A. M. N., Tanvir N., McConnachie A. W., 2006, MN-RAS, 371, 1983

Mayer L., Moore B., Quinn T., Governato F., Stadel J., 2002, MN-RAS, 336, 119

Metcalf R. B., Madau P., 2001, ApJ, 563, 9

Sakamoto T., Hasegawa T., 2006, ApJL, 653, L29

Schechter P. L., Wambsganss J., 2002, ApJ, 580, 685

Springel V., Wang J., Vogelsberger M., et al., 2008, MNRAS, 391,

1685

Stoehr F., White S. D. M., Springel V., Tormen G., Yoshida N., 2003, MNRAS, 345, 1313

Treu T., Gavazzi R., Gorecki A., Marshall P. J., Koopmans L. V. E., Bolton A. S., Moustakas L. A., Burles S., 2009, ApJ, 690, 670

Vegetti S., Koopmans L. V. E., 2009, MNRAS, 392, 945

Walsh S. M., Jerjen H., Willman B., 2007, ApJL, 662, L83

Willman B., Dalcanton J. J., Martinez-Delgado D., et al., 2005, ApJL, 626, L85

Zucker D. B., Belokurov V., Evans N. W., et al., 2006a, ApJL, 650, L41

Zucker D. B., Belokurov V., Evans N. W., et al., 2006b, ApJL, 643, L103

Zucker D. B., Kniazev A. Y., Bell E. F., et al., 2004, ApJL, 612, L121

Zucker D. B., Kniazev A. Y., Martínez-Delgado et al., 2007, ApJL, 659, L21

Table 1. Results for systems with no mass measurement error for a uniform (left) and a Gaussian (right) prior on α. The input parameters are given in columns (1) to (3): the true mass fraction in substructure, f_{true}, the lower detection mass threshold, M_{low}, and the number of lens systems in a given sample, n_l. The maximum posterior values of f and α are given in columns (4) and (5). Columns from (6) to (8) and (9) to (11) list the median, the 68% and 95% CL for f and α respectively.

f_{true} (%)	M_{low} $(10^8 M_\odot)$	n_l	f_{MP} (%)	α_{MP}	f_{med} (%)	$\sigma_{f,68}$ (%)	$\sigma_{f,95}$ (%)	α_{med}	$\sigma_{\alpha,68}$	$\sigma_{\alpha,95}$
0.1	0.3	10	0.14 | 0.14	1.91 | 1.90	0.18 | 0.15	$^{-0.06}_{+0.11}$ | $^{-0.04}_{+0.06}$	$^{-0.09}_{+0.22}$ | $^{-0.07}_{+0.10}$	1.88 | 1.90	$^{-0.37}_{+0.48}$ | $^{-0.05}_{+0.05}$	$^{-0.58}_{+0.81}$ | $^{-0.09}_{+0.09}$
		30	0.11 | 0.10	1.69 | 1.90	0.13 | 0.10	$^{-0.04}_{+0.08}$ | $^{-0.02}_{+0.03}$	$^{-0.07}_{+0.15}$ | $^{-0.04}_{+0.05}$	1.64 | 1.90	$^{-0.25}_{+0.29}$ | $^{-0.03}_{+0.03}$	$^{-0.40}_{+0.51}$ | $^{-0.06}_{+0.05}$
	1.0	10	0.55 | 0.16	2.70 | 1.90	0.60 | 0.20	$^{-0.35}_{+0.70}$ | $^{-0.10}_{+0.10}$	$^{-0.45}_{+1.40}$ | $^{-0.10}_{+0.20}$	2.69 | 1.91	$^{-0.57}_{+0.23}$ | $^{-0.07}_{+0.07}$	$^{-1.10}_{+0.29}$ | $^{-0.11}_{+0.11}$
		30	0.10 | 0.12	1.81 | 1.90	0.20 | 0.13	$^{-0.09}_{+0.30}$ | $^{-0.05}_{+0.06}$	$^{-0.12}_{+0.69}$ | $^{-0.07}_{+0.11}$	2.14 | 1.90	$^{-0.65}_{+0.60}$ | $^{-0.04}_{+0.04}$	$^{-0.93}_{+0.78}$ | $^{-0.07}_{+0.07}$
	3.0	10	1.53 | 0.10	2.77 | 1.90	2.15 | 0.25	$^{-1.70}_{+4.06}$ | $^{-0.15}_{+0.30}$	$^{-2.00}_{+8.32}$ | $^{-0.20}_{+0.55}$	2.78 | 1.92	$^{-0.46}_{+0.16}$ | $^{-0.10}_{+0.10}$	$^{-1.10}_{+0.20}$ | $^{-0.16}_{+0.16}$
		30	0.80 | 0.13	2.62 | 1.90	1.58 | 0.18	$^{-1.19}_{+2.41}$ | $^{-0.09}_{+0.13}$	$^{-1.44}_{+4.67}$ | $^{-0.13}_{+0.27}$	2.78 | 1.91	$^{-0.43}_{+0.16}$ | $^{-0.07}_{+0.07}$	$^{-1.01}_{+0.20}$ | $^{-0.11}_{+0.12}$
0.5	0.3	10	0.45 | 0.50	1.78 | 1.89	0.49 | 0.42	$^{-0.12}_{+0.18}$ | $^{-0.08}_{+0.09}$	$^{-0.18}_{+0.34}$ | $^{-0.13}_{+0.16}$	1.76 | 1.90	$^{-0.19}_{+0.21}$ | $^{-0.03}_{+0.03}$	$^{-0.31}_{+0.37}$ | $^{-0.06}_{+0.05}$
		30	0.49 | 0.50	1.90 | 1.90	0.51 | 0.50	$^{-0.07}_{+0.09}$ | $^{-0.06}_{+0.06}$	$^{-0.11}_{+0.16}$ | $^{-0.09}_{+0.11}$	1.90 | 1.90	$^{-0.12}_{+0.13}$ | $^{-0.02}_{+0.01}$	$^{-0.20}_{+0.23}$ | $^{-0.03}_{+0.03}$
		200	0.51 | 0.50	1.92 | 1.90	0.52 | 0.52	$^{-0.05}_{+0.06}$ | $^{-0.04}_{+0.04}$	$^{-0.08}_{+0.10}$ | $^{-0.07}_{+0.08}$	1.92 | 1.90	$^{-0.09}_{+0.09}$ | $^{-0.01}_{+0.01}$	$^{-0.15}_{+0.15}$ | $^{-0.02}_{+0.02}$
	1.0	10	0.52 | 0.48	2.04 | 1.90	0.85 | 0.52	$^{-0.36}_{+1.28}$ | $^{-0.15}_{+0.18}$	$^{-0.49}_{+2.55}$ | $^{-0.23}_{+0.32}$	2.33 | 1.90	$^{-0.53}_{+0.44}$ | $^{-0.04}_{+0.04}$	$^{-0.83}_{+0.59}$ | $^{-0.06}_{+0.06}$
		30	0.42 | 0.50	1.68 | 1.90	0.47 | 0.44	$^{-0.11}_{+0.14}$ | $^{-0.10}_{+0.11}$	$^{-0.16}_{+0.26}$ | $^{-0.15}_{+0.19}$	1.69 | 1.90	$^{-0.26}_{+0.29}$ | $^{-0.02}_{+0.02}$	$^{-0.43}_{+0.51}$ | $^{-0.04}_{+0.04}$
		200	0.46 | 0.48	1.82 | 1.90	0.49 | 0.50	$^{-0.07}_{+0.10}$ | $^{-0.08}_{+0.09}$	$^{-0.12}_{+0.18}$ | $^{-0.13}_{+0.16}$	1.84 | 1.90	$^{-0.19}_{+0.20}$ | $^{-0.02}_{+0.02}$	$^{-0.30}_{+0.34}$ | $^{-0.03}_{+0.03}$
	3.0	10	1.85 | 0.46	2.56 | 1.90	5.91 | 0.60	$^{-4.01}_{+6.92}$ | $^{-0.30}_{+0.30}$	$^{-5.21}_{+1.28}$ | $^{-0.40}_{+0.60}$	2.82 | 1.91	$^{-0.31}_{+0.13}$ | $^{-0.05}_{+0.06}$	$^{-0.70}_{+0.16}$ | $^{-0.09}_{+0.09}$
		30	0.42 | 0.50	1.68 | 1.90	0.47 | 0.44	$^{-0.11}_{+0.14}$ | $^{-0.10}_{+0.11}$	$^{-0.16}_{+0.26}$ | $^{-0.15}_{+0.19}$	1.69 | 1.90	$^{-0.26}_{+0.29}$ | $^{-0.02}_{+0.02}$	$^{-0.43}_{+0.51}$ | $^{-0.04}_{+0.04}$
		200	0.50 | 0.50	1.90 | 1.90	0.35 | 0.43	$^{-0.09}_{+0.13}$ | $^{-0.12}_{+0.16}$	$^{-0.13}_{+0.30}$ | $^{-0.19}_{+0.28}$	1.35 | 1.90	$^{-0.24}_{+0.38}$ | $^{-0.03}_{+0.03}$	$^{-0.32}_{+0.70}$ | $^{-0.06}_{+0.05}$
2.5	0.3	10	2.50 | 2.56	1.88 | 1.89	2.68 | 2.57	$^{-0.27}_{+0.33}$ | $^{-0.21}_{+0.22}$	$^{-0.44}_{+0.59}$ | $^{-0.34}_{+0.39}$	1.85 | 1.89	$^{-0.08}_{+0.08}$ | $^{-0.03}_{+0.03}$	$^{-0.13}_{+0.14}$ | $^{-0.04}_{+0.05}$
		30	2.39 | 2.31	1.82 | 1.89	2.42 | 2.33	$^{-0.19}_{+0.21}$ | $^{-0.16}_{+0.16}$	$^{-0.31}_{+0.36}$ | $^{-0.27}_{+0.28}$	1.82 | 1.90	$^{-0.07}_{+0.07}$ | $^{-0.02}_{+0.01}$	$^{-0.12}_{+0.13}$ | $^{-0.03}_{+0.03}$
	1.0	10	2.42 | 2.46	1.75 | 1.89	2.54 | 2.50	$^{-0.36}_{+0.42}$ | $^{-0.34}_{+0.39}$	$^{-0.57}_{+0.74}$ | $^{-0.56}_{+0.66}$	1.76 | 1.90	$^{-0.16}_{+0.17}$ | $^{-0.03}_{+0.03}$	$^{-0.27}_{+0.30}$ | $^{-0.05}_{+0.05}$
		30	2.12 | 2.14	1.70 | 1.89	2.16 | 2.15	$^{-0.22}_{+0.25}$ | $^{-0.21}_{+0.23}$	$^{-0.35}_{+0.43}$ | $^{-0.35}_{+0.39}$	1.70 | 1.90	$^{-0.11}_{+0.12}$ | $^{-0.02}_{+0.01}$	$^{-0.18}_{+0.20}$ | $^{-0.03}_{+0.03}$
	3.0	10	2.74 | 2.76	1.89 | 1.90	5.26 | 2.88	$^{-2.76}_{+14.3}$ | $^{-0.63}_{+0.75}$	$^{-3.38}_{+35.8}$ | $^{-1.00}_{+1.25}$	2.22 | 1.90	$^{-0.46}_{+0.44}$ | $^{-0.03}_{+0.03}$	$^{-0.73}_{+0.65}$ | $^{-0.05}_{+0.05}$
		30	1.86 | 2.48	1.28 | 1.89	1.96 | 2.53	$^{-0.32}_{+0.32}$ | $^{-0.40}_{+0.44}$	$^{-0.48}_{+0.60}$ | $^{-0.60}_{+0.76}$	1.31 | 1.90	$^{-0.19}_{+0.21}$ | $^{-0.02}_{+0.02}$	$^{-0.27}_{+0.36}$ | $^{-0.03}_{+0.03}$

APPENDIX A:

We show here that the procedure of first drawing objects between M_{min} and M_{max} from $dP/dm|_{true}$, then scattering with a Gaussian and finally restricting to those masses between M_{low} and M_{high} gives a probability $P(N_{\text{obs}})$ of observing N_{obs} objects that is equivalent to the Poisson probability density distribution, with expectation value μ_s, of N_{obs} objects between M_{low} and M_{high} expressed by the convolution in equation (7).

Let us divide the mass ranges $[M_{\text{min}}, M_{\text{max}}]$ and $\left[M_{\text{low}}, M_{\text{high}}\right]$ re-

Table 2. Results for systems with a mass measurement error of $0.1 \times 10^8 M_\odot$, $0.3 \times 10^8 M_\odot$, and $1.0 \times 10^8 M_\odot$ for a uniform (left) and a Gaussian (right) prior on α. The input parameters are given in columns (1) to (3): the true mass fraction in substructure, f_{true}, the lower detection mass threshold, M_{low}, and the number of lens systems in a given sample, n_l. The maximum posterior values of f and α are given in columns (4) and (5). Columns from (6) to (8) and (9) to (11) list the median, the 68% and 95% CL for f and α respectively.

f_{true} (%)	M_{low} $(10^8 M_\odot)$	n_l	f_{MP} (%)	α_{MP}	f_{med} (%)	$\sigma_{f,68}$ (%)	$\sigma_{f,95}$ (%)	α_{med}	$\sigma_{\alpha,68}$	$\sigma_{\alpha,95}$
0.5	0.3	10	0.53 | 0.61	2.24 | 1.91	0.57 | 0.64	$^{-0.09}_{+0.09}$ | $^{-0.12}_{+0.09}$	$^{-0.15}_{+0.18}$ | $^{-0.15}_{+0.18}$	2.25 | 1.91	$^{-0.20}_{+0.22}$ | $^{-0.02}_{+0.04}$	$^{-0.32}_{+0.36}$ | $^{-0.04}_{+0.06}$
		30	0.53 | 0.53	1.91 | 1.90	0.54 | 0.54	$^{-0.09}_{+0.09}$ | $^{-0.06}_{+0.06}$	$^{-0.12}_{+0.15}$ | $^{-0.09}_{+0.09}$	1.91 | 1.91	$^{-0.12}_{+0.12}$ | $^{-0.01}_{+0.01}$	$^{-0.18}_{+0.20}$ | $^{-0.04}_{+0.01}$
		200	0.51 | 0.52	1.93 | 1.90	0.51 | 0.51	$^{-0.06}_{+0.06}$ | $^{-0.03}_{+0.06}$	$^{-0.09}_{+0.12}$ | $^{-0.06}_{+0.09}$	1.93 | 1.91	$^{-0.08}_{+0.08}$ | $^{-0.01}_{+0.01}$	$^{-0.14}_{+0.14}$ | $^{-0.02}_{+0.01}$
	1.0	10	0.53 | 0.44	2.18 | 1.90	0.82 | 0.48	$^{-0.36}_{+0.94}$ | $^{-0.15}_{+0.15}$	$^{-0.48}_{+1.79}$ | $^{-0.21}_{+0.27}$	2.41 | 1.91	$^{-0.46}_{+0.38}$ | $^{-0.04}_{+0.04}$	$^{-0.77}_{+0.52}$ | $^{-0.08}_{+0.06}$
		30	0.55 | 0.54	1.98 | 1.90	0.64 | 0.54	$^{-0.15}_{+0.18}$ | $^{-0.09}_{+0.12}$	$^{-0.21}_{+0.48}$ | $^{-0.15}_{+0.21}$	2.03 | 1.91	$^{-0.24}_{+0.26}$ | $^{-0.04}_{+0.02}$	$^{-0.40}_{+0.44}$ | $^{-0.04}_{+0.04}$
		200	0.67 | 0.59	2.12 | 1.90	0.72 | 0.61	$^{-0.15}_{+0.18}$ | $^{-0.09}_{+0.09}$	$^{-0.21}_{+0.39}$ | $^{-0.12}_{+0.15}$	2.15 | 1.91	$^{-0.16}_{+0.18}$ | $^{-0.01}_{+0.01}$	$^{-0.28}_{+0.30}$ | $^{-0.04}_{+0.02}$
	3.0	10	4.69 | 0.42	2.85 | 1.90	4.01 | 0.50	$^{-2.76}_{+4.76}$ | $^{-0.25}_{+0.25}$	$^{-3.51}_{+9.02}$ | $^{-0.25}_{+0.50}$	2.78 | 1.91	$^{-0.38}_{+0.16}$ | $^{-0.05}_{+0.06}$	$^{-0.83}_{+0.20}$ | $^{-0.09}_{+0.09}$
		30	2.76 | 0.51	2.57 | 1.90	5.26 | 0.75	$^{-3.51}_{+4.76}$ | $^{-0.25}_{+0.1e-4}$	$^{-4.26}_{+8.02}$ | $^{-0.25}_{+0.25}$	2.76 | 1.91	$^{-0.32}_{+0.17}$ | $^{-0.04}_{+0.03}$	$^{-0.62}_{+0.21}$ | $^{-0.06}_{+0.06}$
		200	0.51 | 0.51	1.90 | 1.90	3.76 | 0.75	$^{-2.25}_{+4.26}$ | $^{-0.25}_{+0.1e-4}$	$^{-3.01}_{+7.02}$ | $^{-0.25}_{+0.25}$	2.68 | 1.90	$^{-0.33}_{+0.22}$ | $^{-0.03}_{+0.03}$	$^{-0.58}_{+0.28}$ | $^{-0.04}_{+0.05}$

spectively in n and n' sub-intervals of infinitesimally small widths dm and dm'. Lets call $p_{s,j}$ the probability that an object is scattered from the j-th mass bin dm_j to the k-th one dm'_k. In the particular case of Gaussian errors, $p_{s,j,k}$ reads as follows

$$p_{s,j,k} = \frac{e^{-(m-m')^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \, dm'_k , \tag{A1}$$

with $m \in dm_j$ and $m' \in dm'_k$.

First we show that, if substructures are Poisson distributed in each mass bin with expectation value $d\mu_j$, then the probability of having N_s objects scattered from dm_j into dm'_k is also a Poisson distribution with expectation value $d\mu_{s,j,k} = p_{s,j,k} \, d\mu_j$

$$P_{s,j}(N_s, d\mu_{s,j,k}) = \sum_{i=N_s}^{\infty} P(i, d\mu_j) \binom{i}{N_s} p_{s,j,k}^{N_s} \left(1 - p_{s,j,k}\right)^{i-N_s} =$$

$$\sum_{i=N_s}^{\infty} \frac{e^{-d\mu_j} d\mu_j^i}{i!} \binom{i}{N_s} p_{s,j,k}^{N_s} \left(1 - p_{s,j,k}\right)^{i-N_s} =$$

$$\frac{p_{s,j,k}^{N_s}}{N_s!} \lim_{n\to\infty} \sum_{i=N_s}^{n} \frac{e^{-d\mu_j} d\mu_j^i}{(i-N_s)!} \left(1 - p_{s,j,k}\right)^{i-N_s} =$$

$$\left(\frac{p_{s,j,k}}{1 - p_{s,j,k}}\right)^{N_s} \frac{e^{-d\mu_j} d\mu_j^{N_s}}{N_s!} \lim_{n\to\infty} \sum_{k=0}^{n-N_s} \frac{d\mu_j^k}{k!} \left(1 - p_{s,j,k}\right)^k =$$

$$\frac{e^{-p_{s,j,k} \mu_j} \left(p_{s,j,k} \, d\mu_j\right)^{N_s}}{N_s!} , \tag{A2}$$

hence it follows that $d\mu_{s,j,k} = p_{s,j,k} \, d\mu_j$. This result directly follows from the fact that in the case of high number statistics the Binomial tends to a Poisson distribution and from the product rule of the Poisson distribution. Specifically, each $d\mu_j$ reads as

$$d\mu_j = \mu_0(\alpha, f, R) \left.\frac{dP}{dm}\right|_{true} dm_j . \tag{A3}$$

We now extend this result to two mass intervals of the same size dm; thanks to the sum rule, the probability of N_{obs} objects being

scattered is again Poissonian with $d\mu_{s,k} = (p_{s,1,k} d\mu_1 + p_{s,2,k} d\mu_2)$

$$P(N_{obs}, d\mu_s) = \sum_{i=0}^{N_{obs}} P_{s,1}(i) P_{s,2}(N_{obs} - i) =$$

$$\sum_{i=0}^{N_{obs}} \frac{e^{-p_{s,1,k} d\mu_1} \left(p_{s,1,k} \, d\mu_1\right)^i}{i!} \frac{e^{-p_{s,2,k} d\mu_2} \left(p_{s,2,k} \, d\mu_2\right)^{N_{obs}-i}}{(N_{obs}-i)!} =$$

$$e^{\left(p_{s,1,k} d\mu_1 + p_{s,2,k} d\mu_2\right)} \left(p_{s,2} \, d\mu_2\right)^{N_{obs}} \sum_{i=0}^{N_{obs}} \frac{d\mu_1^i}{i!} \frac{d\mu_2^{-i}}{(N_{obs}-i)!} =$$

$$e^{-\left(p_{s,1,k} d\mu_1 + p_{s,2,k} d\mu_2\right)} \left(p_{s,2,k} \, d\mu_2\right)^{N_{obs}} \frac{\left(1 + \frac{p_{s,1,k} \, d\mu_1}{p_{s,2,k} \, d\mu_2}\right)^{N_{obs}}}{N_{obs}!} = \frac{e^{-d\mu_{s,k}} d\mu_{s,k}^{N_{obs}}}{N_{obs}!} . \tag{A4}$$

By induction it can be shown that in the case of a generic number n of intervals $d\mu_{s,k} = \sum_{j=1}^{n} p_{s,j,k} \, d\mu_j$, so that the probability of being scattered outside $[M_{\text{min}}, M_{\text{max}}]$ and inside $\left[M_{\text{low}}, M_{\text{high}}\right]$ is a Poisson distribution with expectation value

$$\mu_s = \sum_{k=1}^{n'} \sum_{j=1}^{n} p_{s,j,k} \, d\mu_j = \sum_{k=1}^{n'} \sum_{j=1}^{n} p_{s,j,k} \, \mu_0(\alpha, f, R) \left.\frac{dP}{dm}\right|_{true} dm_j \, dm'_k . \tag{A5}$$

In the limit of equally infinitesimal intervals, i.e $dm \to 0$ and $dm' \to 0$ and making use of equation (A1) this becomes

$$\mu_s = \mu_0(\alpha, f, R) \int_{M_{\text{low}}}^{M_{\text{high}}} \int_{M_{\text{min}}}^{M_{\text{max}}} \frac{e^{-(m-m')^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \left.\frac{dP}{dm}\right|_{true} dm \, dm' . \tag{A6}$$

q.e.d